L. Mingari[1] | A. Folch[1] | A. Guerrero[1] | S. Barsotti[2] | T. Barnie[2] | G. Macedonio[3] | A. Costa[4]

1. Geociencias Barcelona (GEO3BCN-CSIC), Barcelona, Spain
2. Icelandic Meteorological Office (IMO), Iceland
3. Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Naples, Italy
4. Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy

lmingari@geo3bcn.csic.es

SUMMARY

A Digital Twin Component (DTC) provides users with digital replicas of different components of the Earth system through unified frameworks integrating real-time observations and state-of-the-art numerical models. Scenarios of extreme events for natural hazards can be studied from the genesis to propagation and impacts using DTCs. The EU DT-GEO project (2022-2025) is implementing a prototype digital twin on geophysical extremes consisting of 12 interrelated Digital Twin Components, intended as self-contained and containerised software entities embedding numerical model codes, management of real-time data streams and data assimilation methodologies. Single or multiple coupled DTCs can be deployed and executed in centralised High Performance Computing (HPC) and cloud computing Research Infrastructures (RIs). In particular, the DTC-V2 is implementing an ensemble-based automated operational system for deterministic and probabilistic forecast of long-range ash dispersal and local-scale tephra fallout. The system continuously screens different ground-based and satellite-based data sources and a workflow is automatically triggered by a volcanic eruption to stream and pre-process data, its ingestion into the FALL3D dispersal model, a centralised or distributed HPC model execution, and the post-processing step.

METEOROLOGICAL DATASET

FALL3D reads multiple meteorological datasets. To this purpose input from meteorological models are interpolated into the computational grid in the pre-proccessing step. The meteorological models supported are:

- Global Forecast System (GFS) data
- Global Ensemble Forecast System (GEFS)
- Atmospheric reanalyses ERA5 (model levels and pressure levels)
- WRF-ARW: mesoscale model
- Integrated Forecasting System (IFS) data from ECMWF (work in progress).

ENSEMBLE MODELLING

Ensemble simulations can be performed very efficiently by means of a single parallel task. Different ensemble members are built from a reference run by perturbing relevant variables (e.g. column height, wind components, etc.).

- Efficient generation of multiple model outcomes using a single run
- Characterise and quantify uncertainties due to poorly constrained input parameters and model physics parameterisations
- Provide probabilistic products such as the probability of cloud column mass above satellite detection threshold or ash fallout exceeding certain thickness
- Improve forecasts by incorporating observations using different ensemble-based data assimilation techniques

FALL3D

NUMERICAL MODEL: FALL3D

FALL3D is an open-source offline Eulerian model for atmospheric passive transport and deposition of aerosols and particles, including volcanic tephra, volcanic gases, mineral dust, and radionuclides. FALL3D code is available through public GitLab repositories and Zenodo (DOI: 10.5281/zenodo.6343786). This app is developed and maintained by the Geosciences Barcelona and INGV

- FALL3D solves the so-called Advection-Diffusion-Sedimentation (ADS) equation (Folch et al., 2020). This equation is solved on a structured grid using a second order finite volume explicit scheme (Euler or Runge-Kutta 4th scheme in time)
- The new version supports larger scientific workloads and includes several improvements (model physics, numerical algorithmic methods, and computational efficiency)
- From version v8.1 onwards, ensemble simulations can be performed very efficiently by means of a single parallel task

Technical features
- MPI/OpenACC in Fortran 2008
- Ensemble runs using a hierarchy of MPI communicators
- Parallel I/O in NetCDF-4
- Level 2 code: able to perform full capability runs or capacity workloads on current petascale machines and good scalability (>90% efficiency) in about 100 nodes

Code repositories
The FALL3D main repository has already received the gold badge from the EOSC-Synergy SQAaaS. This is a necessary requirement before starting the CI/CD in EuroHPC systems.

A collection of public repositories (suite) exists in GitLab for:

- Main code (CPU version): https://gitlab.com/fall3d-suite/fall3d
- GPU code (GPU version): https://gitlab.com/fall3d-suite/fall3d-gpu
- Mini-app: https://gitlab.com/fall3d-suite/mini-fall3d
- Documentation: https://fall3d-suite.gitlab.io/fall3d/
- Zenodo: https://doi.org/10.5281/zenodo.6343786

Scan me

CODE PERFORMANCE

Strong scaling results have shown perfect scaling with a few hundred processors and parallel efficiencies above 90% up to 2048 processors were found.

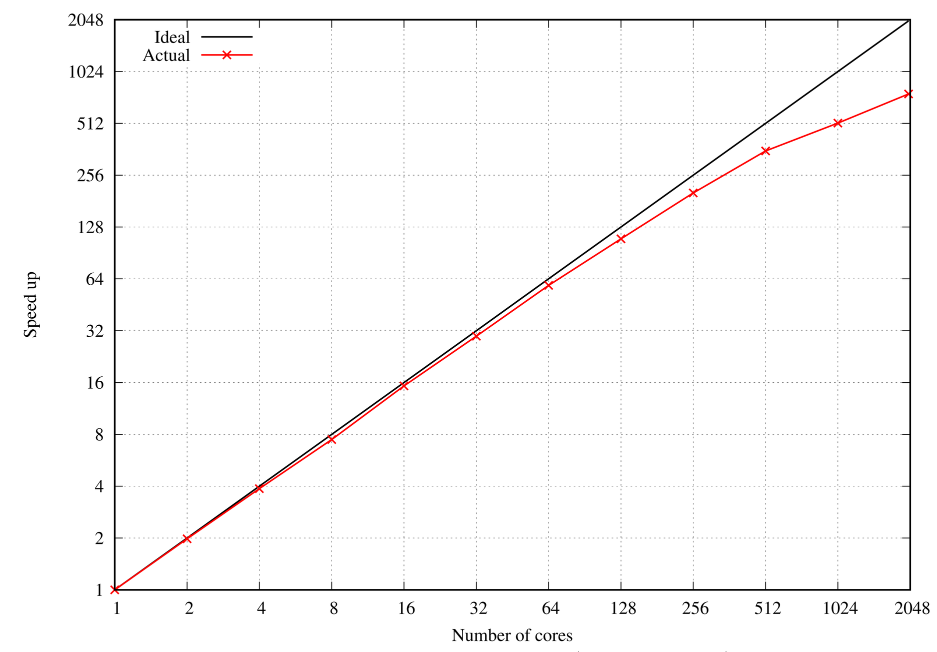

Figure 1: Strong scaling results (speedup) up to 2048 processors obtained on the MareNostrum 4 supercomputer, composed by general-purpose nodes with 48 Intel Xeon Platinum processors interconnected by a 100 GB Intel Omni-Path full-fat tree

NEW FEATURES - WIP

New capabilities are under development in the context of the DT-GEO and ChEESE projects, including:

- A mass-consistent meteorological downscaling in order to improve the meteorological fields over complex terrains. It's based on a diagnostic wind model (e.g. CALMET)
- Implementation of a one-way nesting strategy technique for multiscale simulations
- Design a workflow manager and orchestrator which will enable automated modelling system for atmospheric dispersal and ground deposition of volcanic tephra.

DATA ASSIMILATION

Forecast skills can be substantially improved using an ensemble-based data assimilation system (workflow) combining the FALL3D dispersal model with high-resolution last-generation geostationary satellite retrievals.

Multiple data assimilation methods have been implemented in the FALL3D model, including techniques based on ensemble Kalman filters. New assimilation techniques are currently being explored.

Ensemble-based assimilation methods:
- A new satellite data assimilation module based on the Parallel Data Assimilation Framework (PDAF) has been implemented. PDAF is an open-source software environment for ensemble data assimilation providing fully implemented and optimized data assimilation algorithms.
- The GNC method assumes a multi-dimensional Gaussian distribution and solves an optimization problem with nonnegative constraints to ensure plausible physical solutions and allows non-negative variables, such as aerosol, water vapour, cloud, and precipitation concentrations, to be directly assimilated.

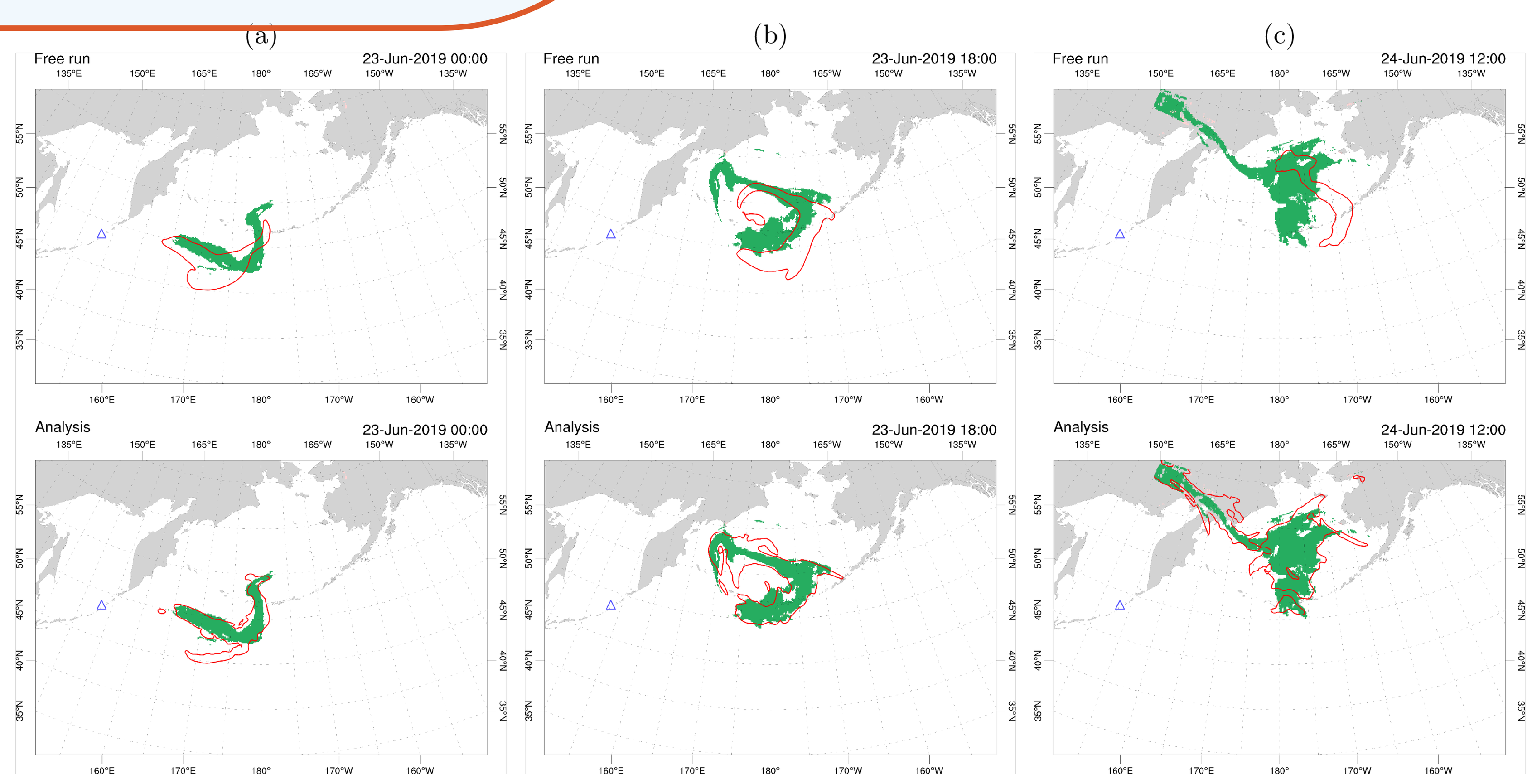

Figure 2: Assimilation of SO2 mass loading from satellite retrievals for the 2019 Raikoke eruption using the FALL3D+PDAF workflow. Model results corresponding to the free run (without assimilation) are shown in the top panel. In the bottom panel, the analyses are shown and compared with satellite detection contours. A dramatic improvement of forecast quality was achieved due to data assimilation.

CONCLUSIONS

The Digital Twin Component for volcanic dispersal and tephra fallout provides frameworks for integrating real-time observations and state-of-the-art numerical models and allows to increase the resolution of existing operational products and improves the forecasting skill by assimilating satellite retrievals. The proposed methodology makes operational ash cloud forecasts compatible with the time-space constraints of aircraft operations which is pivotal in emergency management scenarios and related urgent computing services. Applied over a continental scale, it can be used to produce a 24/7 early warning and forecast system addressed to different stakeholders of the aviation industry, e.g. airlines, Aeronautic Service Providers (ANSPs), airports, local and regional air traffic controllers, etc. On a local/regional scale, it can be used as a 24/7 ash fallout forecast system, addressed to civil protection agencies, decision makers and humanitarian sectors.